

Klasyfikacja danych opisanych za pomocą szeregów czasowych

Promotor: dr inż. Jerzy Błaszczyński

Realizacja: Wojciech Mioduszeński

Cele

▶ Cel

- ▶ Opracowanie i implementacja różnych podejść do klasyfikacji danych czasowych.

▶ Zadania:

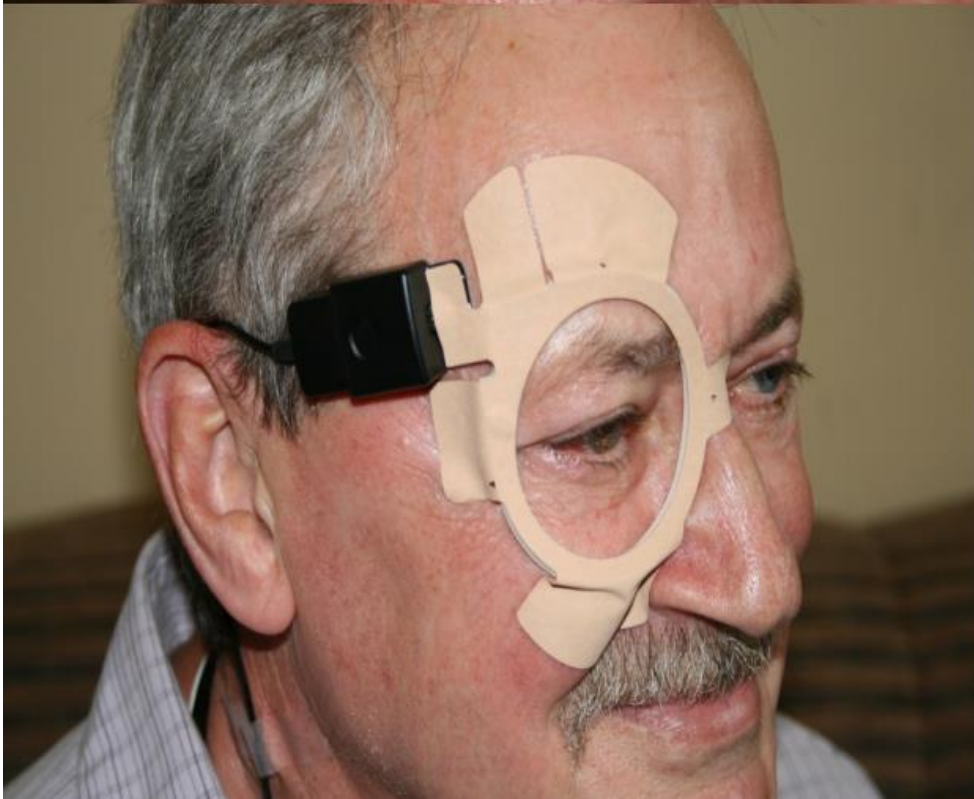
- ▶ Zapoznać się z literaturą tematu.
- ▶ Opracować wybrane podejścia do klasyfikacji danych czasowych.
- ▶ Zaimplementować i udokumentować zaproponowane rozwiązania.
- ▶ Przeprowadzić eksperyment obliczeniowy

Czym jest szereg czasowy

Szereg czasowy to realizacja procesu stochastycznego, którego dziedziną jest czas

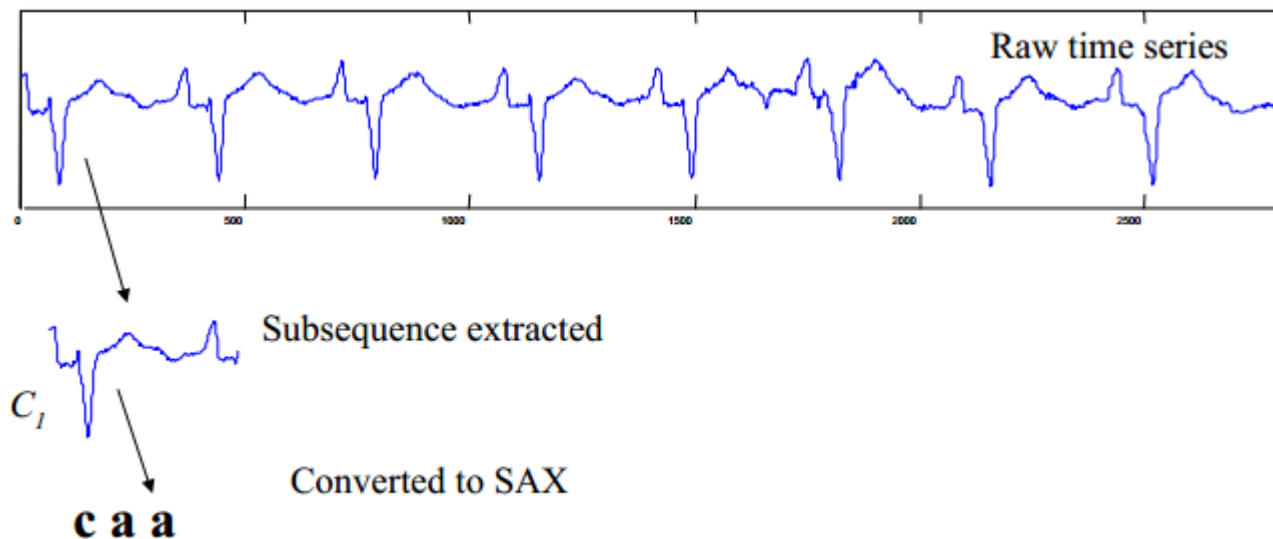
- ▶ Ciąg informacji uporządkowanych w czasie
- ▶ Krok może być regularny, bądź nie - (szeregi czasowe rozmyte)





SAX - Symbolic Aggregate approXimation

- ▶ Eamonn Keogh oraz Jessica Lin (2002 r.)
- ▶ Zamiana szeregu czasowego w ciąg symboli
 - ▶ Redukcja wymiarowości
 - ▶ Standaryzacja długości
 - ▶ Dyskretyzacja



Jak to wykorzystać ?

Dane szeregi czasowe



SAX

Zwykłe dane
(para: ciąg symboli; diagnoza)

Klasyfikator

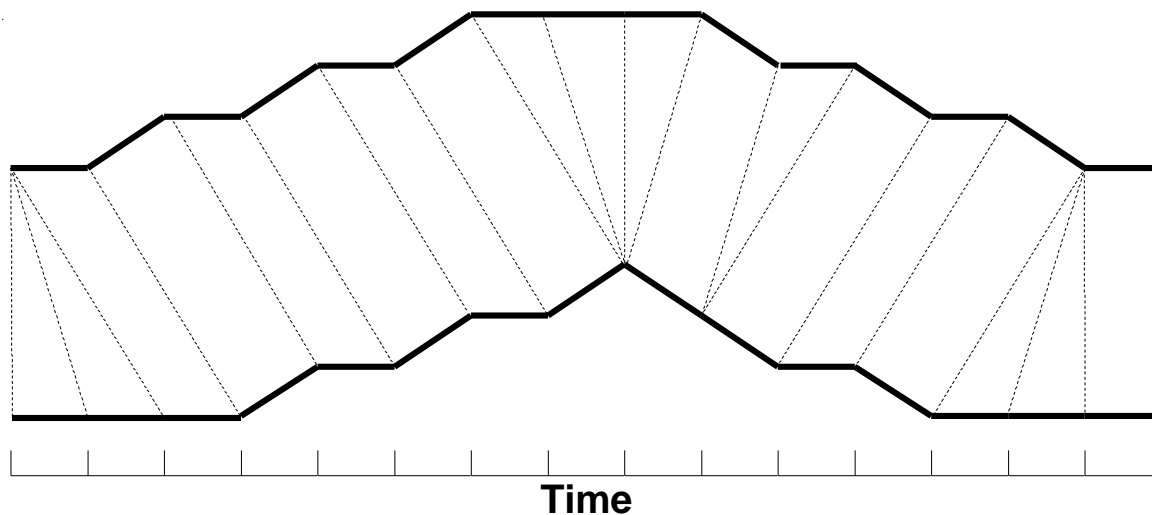


Istniejące klasyfikatory

- ▶ J48
 - ▶ Drzewa decyzyjne
- ▶ JRip
 - ▶ Klasyfikator regułowy
- ▶ Nowy klasyfikator
 - ▶ ???

Dynamic Time Warping

- Podobieństwo między dwoma sekwencjami



- Ratanamahatana C.A., Keogh E.: Everything you know about Dynamic Time Warping is Wrong

Istniejące dane

- ▶ Dane dla jaskry - 20 pacjentów
 - ▶ Długość szeregu czasowego: 96
 - ▶ Dwie klasy (zdrowy, chory)
- ▶ Dane Prof. Eamonna:
 - ▶ 45 problemów
 - ▶ 2-42 klas decyzyjnych
 - ▶ Średnia długość szeregu: 430
 - ▶ Średnia liczba instancji: 650

Dziękuję za uwagę

► Bibliografia:

- Shieh J., Keogh E., iSAX: Indexing and Mining Terabyte Sized Time Series, SIGKDD 2008
- Shumway R., Stoffer D.,: Time Series Analysis and Its Applications, 2011
- Salvador S., Chan P.,: FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space